Examination of Factors Associated with Fault Status and Injury Severity in Intersection-Related Rear-End Crashes: Application of Binary and Bivariate Ordered Probit Models

Extended Summary

Brendan J. Russo[1*], Fan Yu[2], Edward J. Smaglik[3]

1. Corresponding Author, Department of Civil Engineering, Construction Management, and Environmental Engineering, Northern Arizona University, Box 15600 Flagstaff, AZ 86011, USA, brendan.russo@nau.edu*
2. Department of Civil and Environmental Engineering, Northwestern University, 2145 Sheridan Road, Evanston, IL 60208, USA, fan.yu@northwestern.edu
3. Department of Civil Engineering, Construction Management, and Environmental Engineering, Northern Arizona University, Box 15600 Flagstaff, AZ 86011, USA, edward.smaglik@nau.edu

1. Introduction

Rear-end crashes are the most common collision type among police-reported crashes in the United States (US), representing 32.5% of all reported crashes [1]. While rear-end crashes are generally not as severe as some other crash types, there were still 2,346 fatal rear-end crashes and 595,000 injury rear-end crashes (representing 31.1% of all injury crashes) in the US in 2019 alone [1]. Rear-end crashes occur when one vehicle fails to slow or stop and impacts the rear of the vehicle in front of it. These crashes can occur due to variety of reasons (speed, inattentiveness, impairment, weather conditions, etc.), and generally the driver of the following vehicle which makes contact with the lead vehicle is deemed to be at-fault for causing the crash. Given the prevalence of rear-end crashes, it's important to understand factors which may lead to their occurrence and factors which may impact the severity of these crashes when they do occur.

While there is a wealth of previous research generally examining factors associated with traffic crash injury severity outcomes [2] and some research related to the severity of rear-end crash severity does exist [3,4,5,6], there is still knowledge to be gained with respect to these crash types. First, most previous studies investigate rear-end crashes specifically in work zones or across an entire network. While these studies provide important information, none are focused specifically on intersection-related rear end crashes, the characteristics of which may vary compared with those occurring on roadway segments given the variation in speeds and behavior related to interaction with traffic control and possible queuing at intersections. Second, previous work has not investigated factors associated drivers being 'at-fault' and causing the rear-end crash to occur in the first place. Information related to fault-status may be useful in mitigating driver behaviors that lead to these crashes.

To address these gaps in the literature, this study investigates factors related to fault status and injury severity (and the interrelation between the two) in two-vehicle intersection-related rear-end crashes using five years of crash data from the US state of Arizona. A binary probit model was estimated to assess factors associated with fault status, while a bivariate ordered probit model was estimated to assess factors associated with driver injury severity outcomes by fault status. Importantly, by modelling the injury severity of both crash-involved drivers jointly (e.g. bivariate model), potential within-crash correlation can be accounted for. This study adds to the research literature by providing an analysis specifically of injury severity in intersection-related rear-end crashes (including differences observed between intersection types) as well as consideration of the vehicle type for both units involved in the crashes, and factors associated with drivers being at-fault (i.e. causing) these crashes. Ultimately, the results of this study may assist in development of targeted countermeasures (engineering-, enforcement-, and/or education-related) aimed at reducing crash occurrence (through improvement of at-fault driver behavior) and crash severity if such crashes do occur.

2. Methodology

2.1 Data Description

Statewide Arizona crash data for this study were obtained from ADOT for the years 2014 through 2018. The data set was then filtered to include only those crashes coded with a collision manner of 'Rear-End' and flagged as 'Intersection-Related' (generally defined as the crash resulting from an action related to the movement of a vehicle though an intersection – usually within 150 feet or related to queued traffic at an intersection). Additionally, crashes with the following characteristics were filtered out of the data set:

- Any crash involving any number of units other than two.

[1] * Corresponding author. Tel.: +1-928-523-8094;
E-mail address: brendan.russo@nau.edu

- Any crash involving a bicyclist or pedestrian.
- Any crash where the injury severity for either of the two drivers was unknown.
- Any crash in which both drivers or neither driver were identified as 'at-fault'.

Fault status was determined by identifying which driver was issued a violation in the crash. In the context of a two-vehicle rear-end crash, the at-fault vehicle is the one that struck the leading vehicle from behind. The majority of violations for the at-fault driver in this study were for "speed too fast for conditions", "followed too closely", or "inattention/distraction". Ultimately, the final analysis data set included records for 62,006 two-vehicle intersection-related rear-end crashes and 124,012 crash-involved drivers (two drivers for each crash, one at-fault).

All variables included in the crash reports at the incident-, unit- and person-level were considered for this study. The injury status for each driver involved in a crash are described within the Arizona Crash Report Forms Instruction Manual as one the following five discrete categories [7]: Fatal Injury (K-injury), Suspected Serious Injury (A-Injury), Suspected Minor Injury (B-Injury), Possible Injury (C-Injury), or No Injury (O - Property Damage Only). It should be noted that due to the low number of fatal injuries, serious and fatal (A and K) injuries were combined into one category for this study. Table 1 shows a cross tabulation of the injury severities for at-fault and not at-fault drivers. From examination of Table 1, a similar number of at-fault and not at-fault drivers experienced fatal or serious injuries, but it's clear that higher proportion of not at-fault drivers experienced possible or minor injuries as compared with at-fault drivers.

Table 1: Summary of Injury Severity for Crash-Involved Drivers by Fault Status

Injury Severity of Not-At-Fault Drivers	Injury Severity of At-Fault Drivers				
	No Injury (O)	Possible Injury (C)	Minor Injury (B)	Serious or Fatal Injury (A&K)	Total
No Injury (O)	48,395	1,575	1,050	181	51,201 (82.57%)
Possible Injury (C)	6,912	1,078	247	36	8,273 (13.34%)
Minor Injury (B)	1,733	140	396	13	2,282 (3.68%)
Serious or Fatal Injury (A&K)	188	14	23	25	250 (0.40%)
Total	57,228 (92.29%)	2,807 (4.53%)	1,716 (2.77%)	255 (0.41%)	62,006 (100.0%)

The independent variables considered in this study included variables related to the driver (fault status, age, gender, seatbelt use, alcohol/drug use, and distraction), the unit/vehicle (vehicle type and whether vehicle rolled over), the roadway/environment (type of intersection, light condition, horizontal and vertical road alignment, type of lane where crash occurred, speed limit, and surface condition), and the crash itself (season, day of week, county). Of particular interest was intersection type, as crash characteristics may be impacted by the characteristics of an intersection and this potential impact has not been well-studied in previous analyses of rear-end crashes. It should be noted that type of traffic control at each intersection (i.e. traffic signals, stop signs, etc.) was not available for this dataset. All independent variables considered were re-coded into a series of binary variables (e.g. taking the value of 0 or 1) for modeling purposes.

3. Analysis and Results

3.1 Factors Associated with Fault Status using Binary Probit Model

The results of the binary probit model examining factors associated with a driver being at-fault are shown in Table 2. As discussed previously and shown in Table 1, a higher proportion of not at-fault drivers experienced possible or minor injuries as compared with at-fault drivers, indicating drivers who were not at-fault pay a higher price in terms of injury outcomes even though they did not cause the crash. As such, it is important to identify factors associated with being at-fault to examine the underlying driver- or unit-related factors that lead to these crashes occurring.

Table 2: Results of the Binary Probit Fault Status Model

Variable	Coefficient	Std. Error	P-value
Constant	-0.1516	0.0725	0.0365
Car	0.4133	0.0679	<0.0001
Van	0.5060	0.0709	<0.0001
Pickup	0.5190	0.0684	<0.0001
Truck	0.6243	0.0725	<0.0001

Recreational Vehicle (RV)	0.9832	0.2533	0.0001
Motorcycle	0.3720	0.0849	<0.0001
Distracted	1.4276	0.0132	<0.0001
Age 25-63	-0.5962	0.0098	<0.0001
Age 64 and Above	-0.6036	0.0147	<0.0001
Male	0.1425	0.0082	<0.0001
Belted	-0.1295	0.0265	<0.0001
Alcohol or Drugs	2.0244	0.0681	<0.0001
Restricted Log Likelihood	-80,338.90		
Log Likelihood @ Convergence	-69,260.45		

3.2 Factors Associated with Injury Severity using Bivariate Ordered Probit Model

The results of the bivariate ordered probit model examining factors associated with a driver injury severity are shown in Table 3. Overall, numerous roadway, environmental, vehicle, and person related variables were found to associated with driver injury severity, and there was significant positive correlation between the injury severities of drivers involved in the same crash.

Table 3: Results of the Binary Probit Fault Status Model

Variable	At-Fault Driver			Not-At-Fault Driver		
	Coefficient	Std. Error	P-value	Coefficient	Std. Error	P-value
Constant	-1.1711	0.0748	<0.0001	-1.1532	0.0569	<0.0001
Weekend	-0.0375	0.0195	0.0545	-0.0497	0.0148	0.0008
Daylight	-0.0660	0.0216	0.0023	-	-	-
Maricopa County	-0.1349	0.0195	<0.0001	-0.1004	0.0154	<0.0001
T-Intersection	0.1649	0.0208	<0.0001	0.0483	0.0162	0.0029
Roundabout	-	-	-	-0.3981	0.1292	0.0021
Interchange Intersection	0.1371	0.0435	0.0016	-	-	-
Straight Road Alignment	0.1226	0.0468	0.0088	-	-	-
Left Turn Lane	-0.1743	0.0311	<0.0001	-0.0823	0.0231	0.0004
Right Turn Lane	-0.4367	0.0397	<0.0001	-0.1225	0.0232	<0.0001
Speed 30-50mph	0.3170	0.0432	<0.0001	0.1820	0.0304	<0.0001
Speed 55+ mph	0.3320	0.0522	<0.0001	0.2698	0.0362	<0.0001
Surface Ice	-0.9009	0.4752	0.0580	-0.3405	0.1507	0.0238
Van (Driver)	-0.2179	0.0519	<0.0001	-	-	-
Pickup (Driver)	-0.2956	0.0270	<0.0001	-0.0719	0.0192	0.0002
Truck (Driver)	-0.5378	0.0804	<0.0001	-0.3852	0.0584	<0.0001
Bus (Driver)	-0.6454	0.3830	0.0919	-0.6270	0.1196	<0.0001
Motorcycle (Driver)	1.1280	0.0710	<0.0001	1.0073	0.0775	<0.0001
Van (Other vehicle)	0.1689	0.0439	0.0001	0.0719	0.0356	0.0437
Pickup (Other vehicle)	0.1982	0.0223	<0.0001	0.0753	0.0169	<0.0001
Truck (Other vehicle)	0.4812	0.0507	<0.0001	-	-	-
Bus (Other vehicle)	0.5344	0.0918	<0.0001	0.2875	0.1355	0.0339
Motorcycle (Other vehicle)	-0.2226	0.1230	0.0704	-0.8229	0.1119	<0.0001
Rollover	1.2862	0.1422	<0.0001	1.4337	0.1621	<0.0001
Distracted	0.0856	0.0180	<0.0001	0.0768	0.0338	0.0231
Age 25-63	0.0803	0.0193	<0.0001	0.2173	0.0196	<0.0001
Age 64 and Above	0.2740	0.0291	<0.0001	0.2435	0.0257	<0.0001
Female	0.2156	0.0176	<0.0001	0.2129	0.0135	<0.0001
Belted	-0.8122	0.0354	<0.0001	-0.1911	0.0448	<0.0001
Alcohol & Drugs	0.3334	0.0386	<0.0001	-	-	-
Threshold 1	0.4686	0.0099	<0.0001	0.8033	0.0096	<0.0001
Threshold 2	1.4227	0.0286	<0.0001	1.7530	0.0248	<0.0001
Correlation (ρ)	0.4314	0.00961	<0.0001			
Restricted Log Likelihood	-42,123.29					
Log Likelihood @ Convergence	-41,313.29					

4. Discussion

4.1 Discussion of Fault Status Model Results

Based on the results presented in Table 4, the two strongest predictors of a driver being at-fault are the use of alcohol/drugs and being distracted. These are both expected results as these behaviors are associated with degraded driving capability and may be associated with inherently riskier drivers. With respect to age, drivers age 24 years or younger (left out of the model as a reference category) were most likely to be at-fault in a crash, while older drivers (age 64+) were least likely to be at-fault. This may be a result of both the inexperience and possibly riskier driving behavior or younger drivers. The model also indicates that male drivers are more likely to be at fault than females, but further research may be needed for insights into the reasons for this finding.

Drivers who were belted were less likely to be at-fault compared with belted drivers, again likely due to unbelted drivers being inherently riskier drivers. Finally, with respect to vehicle type, RV drivers were most likely to be at-fault, possibly due to the driver's unfamiliarity with driving such vehicles because they are likely not driven as regularly as other vehicle types. Drivers of busses (left out as the reference category) were least likely to be at-fault, possibly due to the experience and expertise of professional bus drivers who are most likely to be driving these vehicles. Overall, the findings from this model provide insights into the driver- and unit-related characteristics associated with causing rear-end crashes at intersections, and may be useful in planning strategies to reduce the frequency of such crashes.

4.2 Discussion of Driver Injury Severity Results

Based on the results presented in Table 5, numerous variables were found to be significantly associated with the injury severity of both at-fault and not-at-fault drivers (though some variables were significant for only one driver or the other), and there was significant correlation in the injury severities of drivers involved in the same crash as evidenced by the correlation coefficient (ρ). First, crashes occurring on weekends and during daylight hours were found to decrease the probability of severe injury for both drivers, results consistent with past research [5,6]. Crashes occurring in Maricopa county (the largest and most urbanized area in the state) were associated with less severe injuries – likely a surrogate for urban areas vs rural areas where speeds may tend to be higher the density of intersections may be lower. With respect to intersection type, crashes occurring at T-intersections (i.e. 3-leg) were associated with the most severe injury outcomes. Interestingly, roundabouts were associated with decreased injury severity for not-at-fault drivers (the variable was not significant for at-fault drivers), while intersections related to an interchange were found to be associated with more severe injury outcomes only for at-fault drivers. With respect to roadway alignment and lane type, straight roadway alignments (as opposed to curved) were associated with more severe injury outcomes while crashes occurring in left or right turn lanes (as opposed to through lanes) were both associated with more severe injury outcomes – both results likely related to speeds at the time of the crash. In terms of speed limit, crashes occurring on roads with higher speed limits were associated with more severe injury outcomes for both driver types - an expected result given the higher impact forces at higher speeds.

With respect to vehicle type, the potential impact of both the vehicle a driver was in and the other vehicle involved in the crash were investigated. For both driver types, drivers of motorcycles were most likely to experience severe injury outcomes, an expected result consistent with prior research [5]. Conversely, drivers in crashes where the other vehicle involved was a motorcycle had a lower probability of severe injuries, likely due to the relatively lower mass of this vehicle type. Drivers of trucks and buses tended to experience the lowest probability of severe injuries, while drivers in crashes where the other vehicle was a truck (for at-fault drivers) or buses (for both drivers) had a higher probability of severe injuries, likely due to the relatively large mass of these vehicle types.

Turning to person-level driver characteristics, male drivers and drivers using seatbelts both had lower probability of severe injury outcomes, both consistent with prior research [5,6]. With respect to driver age, for both driver types younger drivers (age 24 or less) were least likely to be severely injured, while older drivers (age 64+) were most likely to be severely injured. These results may be due to physiological differences between age groups and are consistent with previous studies [3,6]. Driver distraction was associated with increased probability of severe injuries for both driver types. This variable was generally not considered in previous studies in this area but is an intuitive finding as distracted at-fault drivers are less likely to slow down before impact, while distracted not-at-fault drivers would likely not have a chance to 'brace' for impact if they do not see the at-fault vehicle approaching in their rear-view mirror. Finally, alcohol or drug use by the at-fault driver was associated increased probability of severe injury, but interestingly, alcohol or drug use by the not-at-fault driver was not a significant predictor of injury severity.

5. Conclusions

Rear-end crashes are the most common collision type in the US and result in thousands of fatalities and hundreds of thousands of injuries each year [1]. Given these statistics, this study presented an analysis of factors associated with fault status and injury severity outcomes for two-vehicle intersection-related rear-end crashes in an effort to gain insights which may be useful in mitigating these crashes and resultant injuries. Five years of crash data from the US state of Arizona were utilized in developing a binary probit model to assess factors associated with fault status and a bivariate ordered probit model to assess factors associated with driver injury severity outcomes by fault status. Importantly, by modelling the injury severity of both crash-involved drivers jointly (e.g. bivariate model), potential within-crash correlation can be accounted for. The results showed that numerous variables were significantly associated with fault status and injury severity, and that the injury severity of drivers involved in the same crash were positively correlated. This study adds to the research literature by focusing specifically on intersection-related rear end crashes including consideration of intersection type and the type of vehicle driven by both the at-fault and not-at-fault driver in each crash.

Based on the results of the binary probit fault status model, any enforcement or education related activities aimed at preventing intersection rear-end crashes should focus on drivers of RVs and trucks, young drivers (age 24 or less) and male drivers, all characteristics which were found to be associate with being at-fault in causing these crash types. Additionally, efforts should continue to reduce the prevalence of distracted and impaired driving as these were both also found to be associated with being at-fault. With respect to mitigating severe injury outcomes, based on the results of the bivariate ordered probit model investigating injury severity, engineering efforts should be focused on T-intersections and interchange-related intersections, particularly in rural areas with poor lighting conditions or high speed limits. Potential engineering countermeasures at such locations include improved lighting, warning signs (particularly at unsignalized intersections), high-friction pavement treatments, speed limit reduction, and improved signal visibility and/or clearance interval timing at signalized intersections. In terms of enforcement or education related activities, efforts should be focused on prevention of distracted and impaired driving in general, and specifically on behaviors of truck, bus, and motorcycle drivers.

There were some limitations in this study, most of which can also be considered future research directions for possible extension of this work. First, actual speed data for crash-involved vehicles were not available for this study (speed limit was used as a surrogate), however if such data were available, more accurate insights on the impact of speed (particularly speed of at-fault vehicles) on injury severity could be obtained. Second, this study focused only on driver injury severity, though it's likely passengers are present for many crashes of this type. Future work could investigate factors associated with passenger injury severity, particularly for passengers in the rear seat of not-at-fault vehicles who may be most impacted by this crash type. Finally, data for this study were obtained only from the US state of Arizona. The results are likely transferable to other southwest US regions and similar areas, but may not represent all geographic regions as weather conditions, driver behavior, and other factors may vary. Future work could extend the analysis using crash data from across different regions to investigate how results might vary area to area.

References

1. National Highway Traffic Administration, US Department of Transportation. "Traffic Safety Facts 2019: A Compilation of Motor Vehicle Crash Data". Report DOT HS 813 141, 2021.
2. Savolainen, P.T., Mannering, F.L., Lord, D., and M. A. Quddus, "The statistical analysis of highway crash-injury severities: A review and assessment of methodological alternatives," Accident Analysis and Prevention, vol. 43, no. 5, pp. 1666-1676, 2011.
3. Yu, M., Zheng, C., and C. Ma, "Analysis of injury severity of rear-end crashes in work zones: A random parameters approach with heterogeneity in means and variances". Analytic Methods in Accident Research (27), 2020.
4. Zhang, K., and M. Hassan.. "Investigating Factors Contributing to Injury Severity in Work Zone Rear-End Crashes." Journal of Advanced Transportation, Volume 2019, 201
5. Chen, C., Zhang, G., Tarefder, R., Ma, J., and H. Wei, "A Multinomial Logit Model-Bayesian Network Hybrid Approach for Driver Injury Severity Analyses in Rear-End Crashes." Accident Analysis and Prevention 80: 76–88., 2015.
6. Chen, F., Song, M., and X. Ma., "Investigation on the Injury Severity of Drivers in Rear-End Collisions between Cars Using a Random Parameters Bivariate Ordered Probit Model." International Journal of Environmental Research and Public Health 16(14), 2019.
7. Arizona Department of Transportation (ADOT). Arizona's Crash Report Forms Instruction Manual. 11th Edition, 2017.
8. Washington, S.P., M. G. Karlaftis and F. L. Mannering, Statistical and Econometric Methods for Transportation Data Analysis (2nd ed.), Boca Raton, FL: Chapman and Hall/CRC, 2011.
9. Mannering, F.L., and C. R. Bhat, "Analytic methods in accident research: Methodological frontier and future directions," Analytic Methods in Accident Research, vol. 1, pp. 1-22, 2014.
10. Green, W.H., and Hensher, D.A., "Modeling ordered choices: a primer", 1st. ed., Cambridge University Press, Cambridge, United Kingdom, 2009.
11. Anastasopoulos, P.C., Karaftis, M.G., Haddock, J.E., and Mannering, F.L, "Household automobile and motorcycle ownership analyzed with random parameters bivariate ordered probit model". Transportation Research Record 2279, 12-20, 2012.